Université de Montréal
C.P. 6128, Succursale Centre-ville
Montréal (Québec) Canada H3C 3J7

Téléphone : +1-514-343-7575
Télécopieur (Fax) : +1-514-343-7121
crt@crt.umontreal.ca
www.crt.umontreal.ca/CRT/

CRT-2000-28

Centre de recherche sur les transports (C.R.T.)
Centre for Research on Transportation

ACCURATE CALCULATION OF HAZARDOUS MATERIALS TRANSPORT RISKS

by

Bahar Y. Kara
Erhan Erkut
Vedat Verter

September 2000

101-01761/

CRT-2000-28

*ACCURATE CALCULATION OF HAZARDOUS MATERIALS TRANSPORT RISKS**

by

Bahar Y. Kara[1]
Erhan Erkut[2]
Vedat Verter[3]

* *This research was supported in part by the Fonds pour la formation de chercheurs et l'aide à la recherche (FCAR) and by the Canadian Natural Sciences and Engineering Research Council (NSERC). This support is gratefully acknowledged.*

[1] Faculty of Management, McGill University, Montréal, CANADA H3A 1G5 <yetis@management.mcgill.ca>

[2] Faculty of Business, University of Alberta, Edmonton, Alberta, CANADA T6G 2R6

[3] Faculty of Management, McGill University, Montreal, Canada H3A 1G5 and Centre de recherche sur les transports, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, CANADA H3C 3J7 <verter@management.mcgill.ca>

Dépôt légal – Bibliothèque nationale du Québec,
Bibliothèque nationale du Canada, 2000

Abstract

We propose a path selection algorithm for transport of hazardous materials. This algorithm can deal with link impedances that are path-dependent. This approach is superior to the use of standard shortest path algorithm, common in the literature and practice, which results in inaccuracies.

Keywords: Hazardous material transport, shortest path

Résumé

Nous proposons un algorithme de sélection de voie d'accès pour le transport des marchandises dangereuses. Cet algorithme peut traiter les impédances de lien qui sont dependants de la route. Cette approche est supérieure à l'utilisation de l'algorithme standard algorithme du chemin le plus court, courant dans la littérature et la pratique, qui a comme conséquence les inexactitudes.

Mot clefs: Transport de marchandises dangereuses, chemin le plus court.

1. Introduction

Almost all papers that deal with the selection of a minimum risk path for transport of a hazardous material (hazmat) reduce the path selection problem to a shortest path problem. The use of a standard shortest path algorithm requires that the impedance of each transport link be known and independent of the impedances of the other links. However, this is not the case for hazmat transport problems and simplifications are necessary to make the problem solvable using a standard shortest path algorithm. In this paper we discuss the inaccuracies that result from imposing the standard shortest path model on the hazmat transport problem and we propose modifications to a well-known shortest path algorithm which negate the need to make simplifications.

Consider the problem of selecting a path for a shipment of dangerous goods between a pre-specified origin-destination pair. Let $N = \{1, ..., n\}$ denote the node set of the transportation network. Nodes i and j in N are connected via link (i,j). Let P denote a feasible path for this shipment. For the ease of exposition, we assume that the nodes in P are indexed sequentially, i.e. $P=\{1, 2, 3, ..., r\}$, where node 1 represents the origin and node r represents the destination.

What differentiates hazmat transport models from other transport models is the explicit modeling of transport risk which usually consists of one or both of the following two factors: incident (i.e. spill, fire) probability and population impacted. Let p_{ij} denote the probability of having an incident on link (i,j), and let C_{ij} denote the total number of people who live within a given threshold distance of link (i,j).

Using this notation we list two objectives for selecting a hazmat path that are the popular among academics and practitioners.

1) Minimize total population exposure (for example ReVelle et al., 1991):

$$C_{12} + C_{23} + C_{34} + \ldots + C_{r-1r} \qquad (1)$$

2) Minimize total incident probability (for example Saccomanno and Chan, 1985):

$$p_{12} + (1-p_{12})p_{23} + (1-p_{12})(1-p_{23})p_{34} + \ldots + (1-p_{12})\ldots(1-p_{r-2,r-1})p_{r-1r} \qquad (2)$$

We now discuss the simplifications made when using these objectives, the resulting inaccuracies, and how one could avoid these.

2. Modeling Population Exposure

Consider a hazmat truck moving on a link. The impact area of an incident is usually assumed to be a circle, with a substance-dependent radius, centered at the incident location. Consider the union of all such circles centered at all points on link ***(i,j)***. We refer to this area as the "semicircular exposure zone." Figure 1 shows the semicircular exposure zone of radius λ for link ***(i,j)***.

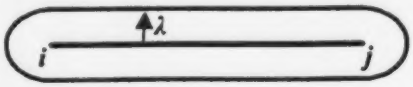

Figure 1: Semicircular exposure zone around link ***(i,j)***

Note that it is possible to preprocess the population density data and compute C_{ij} for all links in the transport network. This, in turn, would allow one to use a standard shortest path algorithm, such as Dijkstra (1959), to find the path that minimizes the population exposure. This is rather tempting since all of this can be accomplished easily using a geographical information system (GIS) such as ArcView (ESRI, 1996). However, such an approach would overestimate the population exposure of all paths in the network to varying degrees and may result in the selection of a suboptimal path.

To demonstrate the overestimation, consider a simple example: a path with only two links, $P = \{1, 2, 3\}$ where the two links intersect with an acute angle. The semicircular exposure zone of this path is given in Figure 2a and the semicircular exposure zones of the two links are shown in Figure 2b.

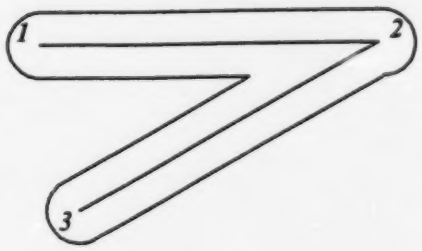

Figure 2a: Semicircular exposure zone around the path

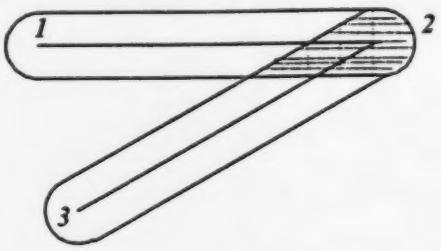

Figure 2b: Semicircular exposure zones of the two links

Observe that the semicircular exposure zone of the path, given in Figure 2a, is not the sum of the semicircular exposure zones of the two links (*1,2*) and (*2,3*), given in Figure 2b. In finding the population exposure of this path, a standard shortest path algorithm would add the population exposure figures of the two links. Clearly, this would double-count the population in the shaded area in Figure 2b resulting in an overestimation.

It is possible to reduce the error by using exposure zones in the shape of rectangles around the links as shown in Figure 3. This is the method used by PC*HazRoute (ALK Associates, 1994), a special software developed for hazmat transportation, to compute population exposure.

Figure 3: Rectangular exposure zone around link (*i,j*)

It is worth noting that if one uses rectangular exposure zones one would be underestimating the exposure at the origin and the destination points due to the lack of the half circles at these points. However, this omission (of a constant) does not impact the optimization problem for path selection, and it can be corrected easily by adding the omitted exposure after finding the optimal path.

More importantly this representation also overestimates the population exposure at intersection points, though not as badly as the semicircular representation. To demonstrate the overestimation, consider the example in Figure 4 where two links, (*1,2*) and (*2,3*) are perpendicular to each other.

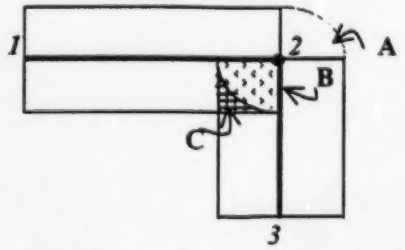

Figure 4: Rectangular exposure zones for a path

With the rectangular exposure zones, the people living in the area marked with A are not counted as being impacted by the transport activity, whereas those living in the area marked with B are counted twice (one for each rectangle). Since population density is assumed to be uniform, the double-counting in B negates the exclusion in A. However, there is still double-counting in C; the area that is double counted is: $\lambda^2 - \pi\lambda^2/4$.

We can express the errors resulting from using a semicircular or rectangular exposure zone at the intersection of two links (k,i) and (i,j) as a function of the population density at the intersection (denoted by $\rho(i)$), the threshold distance for exposure (λ) and the angle between the two links (denoted by α). For the rectangular exposure zone, the error is

$$\varepsilon_{kl,ij} = [\lambda^2/\tan(\alpha/2) - (180-\alpha)\pi\lambda^2/360]\,\rho(i), \qquad (3)$$

and for the semicircular exposure zone the error is equal to the above expression plus the number of people living in the two semi-circles at the intersecting node, namely $\rho(i)\pi\lambda^2$.

As per (3) the increase of the error is linear in ρ and quadratic in λ. As one would expect, the error for the rectangular representation goes to zero as the angle approaches 180 degrees (no double counting if the links are lined up perfectly), and it increases drastically as the angle becomes smaller. Figure 5 shows the errors as a function of α for $\lambda = 1$ and $\rho = 100$ (a 1 km. threshold and a population density of 100 per square-km).

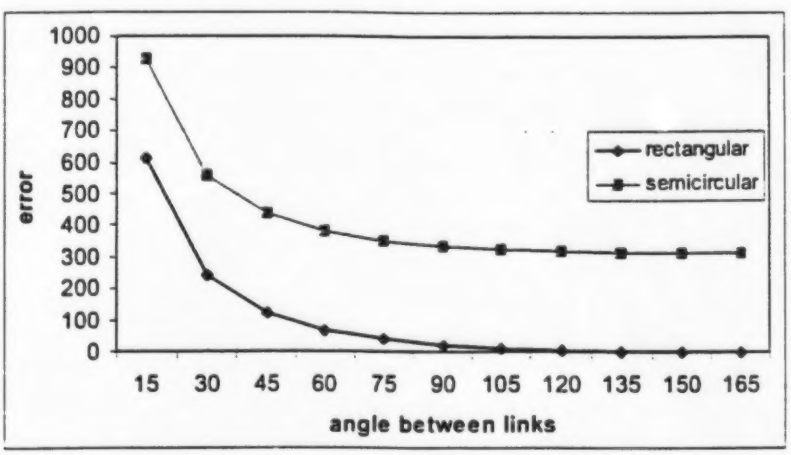

Figure 5: The double-counting error for the semicircular and the rectangular exposure zones as a function of the angle between two adjacent links for a fixed distance threshold ($\lambda = 1$) and fixed population density ($\rho = 100$).

The semicircular representation can lead to fairly significant errors due to the double-counting at the nodes. It is clear that the rectangular representation is associated with lower error figures and in many cases the resulting errors may be negligible. However, the errors associated with rectangular representation can reach nontrivial values depending on the population density around the link intersections and the angle between the link pairs. For example, within Montreal there is a major highway junction point where six links intersect. Using population figures of 1996, the overestimation for a link pair at this junction can be as high as 797 people for gasoline flows ($\lambda = 800$ meters) resulting in a relative error of 1.3%. The intersection of four highway segments at Sainte-Foy, Quebec, offers another example. The overestimation at this junction can reach 157 people for a relative error of 3% with $\lambda = 800$ meters. We provide a complete numerical example in the Appendix that demonstrates the errors resulting from the use of a standard shortest path algorithm with the two population exposure zone representations on the highway network of South-Western Ontario, Canada.

3. An Impedance-Adjusting Shortest Path Algorithm

The error term is a function of three parameters, and we could calculate it for all pairs of adjacent links during preprocessing. This allows us to use initial population exposure as algorithm input and then correct for the double-counting on-the-fly. We now propose a revised version of Dijkstra's shortest path algorithm to find minimum population exposure paths. The algorithm adjusts the link impedances at each iteration, where the adjustment consists of the subtraction of the double-counted portion of the population exposure. The required information for the adjustment at a node is determined by the predecessor node. Although shortest path algorithms for dynamically adjusted link lengths have been proposed in the operational research literature for other problems (for example Dreyfus, 1969, Sung et al., 2000) we know of no reference to them in the hazmat transportation literature.

Now, we present an algorithm, which correctly computes the population exposure of each path, and identifies the path with minimal total population exposure. At a given iteration

of the algorithm, let $\theta(i)$ denote the length of the current shortest path to node i and c_{ij} denote the increase in the objective value due to the addition of link (i,j) to the current path. Let $pred(i)$ denote the predecessor of node i. In standard shortest path formulations c_{ij} is a problem parameter that does not depend on the predecessor node. In our case, with $k = pred(i)$, c_{ij} is computed as:

$$c_{ij} = C_{ij} - \varepsilon_{ki,ij} \tag{4}$$

Note that $C_{ij} = 2\lambda l_{ij}\rho_{ij}$ for rectangular exposure zones and $C_{ij} = 2\lambda l_{ij}\rho_{ij} + 1/2\pi\lambda^2(\rho(i) + \rho(j))$ for semicircular zones, where l_{ij} denotes the link length and ρ_{ij} denotes the population density around the link. Likewise, the calculation of the error term in (4) should be consistent with the type of exposure zone.

Impedance-Adjusting Shortest Path Algorithm:

Initialize $S = \{\}$, $S' = N$, $\theta(i) = \infty\ \forall i$, $\theta(1) = 0$, $pred(1) = 1$, $\varepsilon_{11,\,1j} = 0\ \forall j$ incident to 1.
While $r \notin S$;
  Let i in S' be such that $\theta(i) = min\ \{\theta(j): j \in S'\}$
  $S = S \cup \{i\}$, $S' = S' \setminus \{i\}$
  For each adjacent link (i,j) to node i
    Calculate $\varepsilon_{pred(i)i,\,ij}$ and c_{ij}
    If $\theta(j) > \theta(i) + c_{ij}$ then $\theta(j) = \theta(i) + c_{ij}$ and $pred(j)=i$

The total number of people exposed is $\theta(r)$ for the semicircular zones and $\theta(r) + \pi\lambda^2(\rho_1 + \rho_r)/2$ for rectangular zones.

4. Modeling Incident Probability

Note that (2) computes the probability of an incident along a path by adding the incident probabilities along each link of the path. The probability of incident on a given link depends on probabilities of incident along all links leading up to the given link. Hence, the incident probabilities on links are path-dependent. While it is possible to formulate a nonlinear integer programming problem for selecting the minimum probability path, such a formulation is useless for practical purposes. All papers that deal with (2) use a

simplification which turns this complicated optimization problem to a shortest path problem: products of incidence probabilities can be approximated by zero. This assumption is justified by the magnitude of incident probabilities (usually on the order of 10^{-6} incidents per mile). Erkut and Verter (1998) point out that this approximation is likely to result in a very small error (less than 0.25% in most cases) in measuring the incident probability along a path. While this error is acceptable for all practical purposes, it is rather easy to apply the impedance adjustment idea discussed in Section 2 to produce error-free estimates.

Let $q(t)$ denote the probability of safely arriving at node t of path P. Note that this probability is dependent on the previous links of the path. Observe that

$$q(t+1) = q(t)(1-p_{t,t+1}),$$

and the definition of impedance becomes:

$$c_{ij} = q(i)p_{ij}.$$

The algorithm must compute the necessary $q(i)$ values at each iteration since they depend on the predecessor nodes. The modification to the algorithm given in Section 3 is straight-forward.

5. Concluding Remarks

The impedance-adjusting shortest path approach proposed here would also work for finding paths that minimize the expected number of people that would suffer the consequences of a possible hazmat incident (which is the traditional definition of risk). Denote the number of people who live inside the danger circle for a given substance on link (i,j) as D_{ij}. The expected consequence of a path is computed by

$$D_{12}p_{12} + D_{23}(1-p_{12})p_{23} + \ldots + D_{r-1,r}(1-p_{12})(1-p_{23})\ldots(1-p_{r-2,r-1})p_{r-1,r}$$

Typically, researchers and practitioners leave out terms involving products of probabilities to reduce the path selection problem to a shortest path problem. In this case

it is more difficult to estimate the error that results from this approximation since it is a function of probabilities as well as consequences. However, an impedance-adjusting shortest path algorithm (which computes the probability of reaching node *t* dynamically) makes this approximation unnecessary.

To conclude, we point out that an impedance-adjusting shortest path algorithm can effectively remove errors in quantifying path impedances for the most popular hazmat route selection methods. In the case of population exposure, the errors result from the mistreatment of input data (compounded by the network topology) whereas in the case of incident probabilities the errors result from a simplification in the original model. While both types of errors are rather small (unless the semicircular exposure representation is used to quantify population exposure--which may result in significant errors), there is no reason for accepting such errors since it is very easy to eliminate them.

References:

ALK Associates (1994), *ALK's PC*HazRoute (Version 2.0)*, ALK Associates, Inc., 1000 Herrontown Road, Princeton, NJ 08540, USA.

Dijkstra, E. (1959), "A Note on Two Problems in Connection with Graphs," *Numerical Mathematics* **1**, 269-271.

Dreyfus S.E. (1969), "An Appraisal of Some Shortest-Path Algorithms," *Operations Research* **17**, 395-412.

Erkut, E. and V. Verter (1998), "A Framework for Hazardous Materials Transport Risk Assessment," *Risk Analysis* **15**(5), 589-601.

ESRI (1996), *ArcView 3.1*, ESRI Inc., 380 New York Street, Redlands, CA 92373-8100, USA.

ReVelle, C., J. Cohon and D. Shobrys (1991), "Simultaneous Siting and Routing in the Disposal of Hazardous Wastes," *Transportation Science* **25**(2), 138-145.

Saccomanno, F. F. and A. Chan (1985), "Economic Evaluation of Routing Strategies for Hazardous Road Shipments," *Transportation Research Record* **1020**, 12-18.

Sung, K., M.G.H. Bell, M. Seong, and S. Park (2000), "Shortest Paths in a Network with Time Dependent Flow Speeds," *European Journal of Operational Research* **121**, 32-39.

Appendix

Figure 6 depicts the population centers and the highway network of South-Western Ontario, Canada. As an illustrative example, we focus on the shipments between Halton Hills and Ancaster. In 1998, there were 577 fuel oil trucks, 200 gasoline trucks and 1382 petroleum trucks shipped between these two population centers.

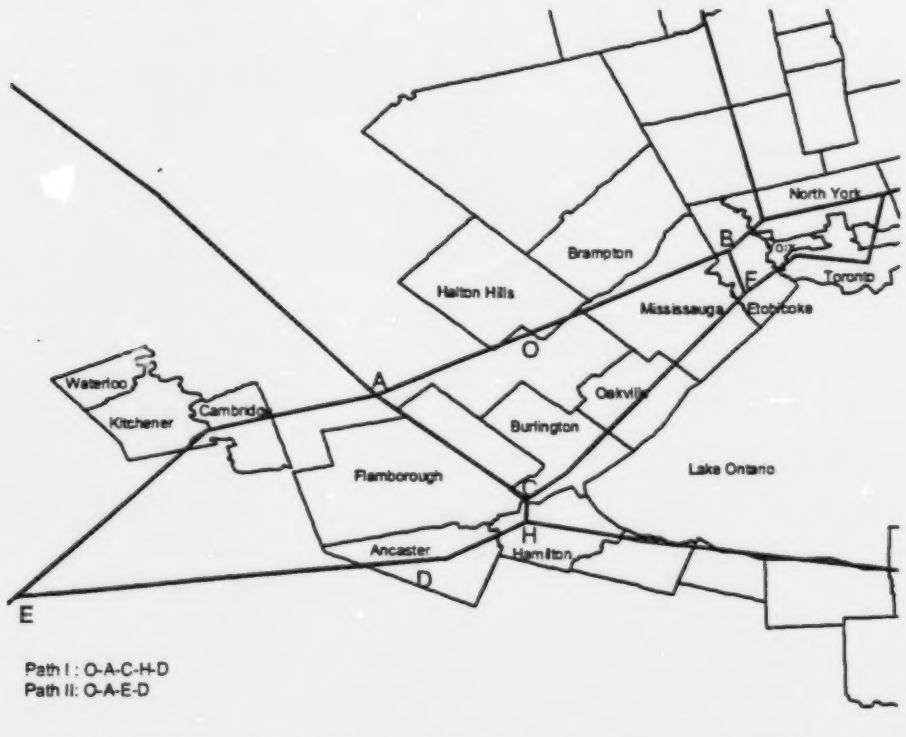

Figure 6: The population centers and highway network of South-Western Ontario

Our analysis is relevant for the gasoline and fuel oil trucks, which would require the use of an 800m exposure zone around the highway route in case of fire. Figure 7 shows the exposure zones in Kitchener and Cambridge, Ontario. Note that there are four highway links in this region, dictated by the shape of the Ontario Highway 401 as well as the boundaries of the two population centers.

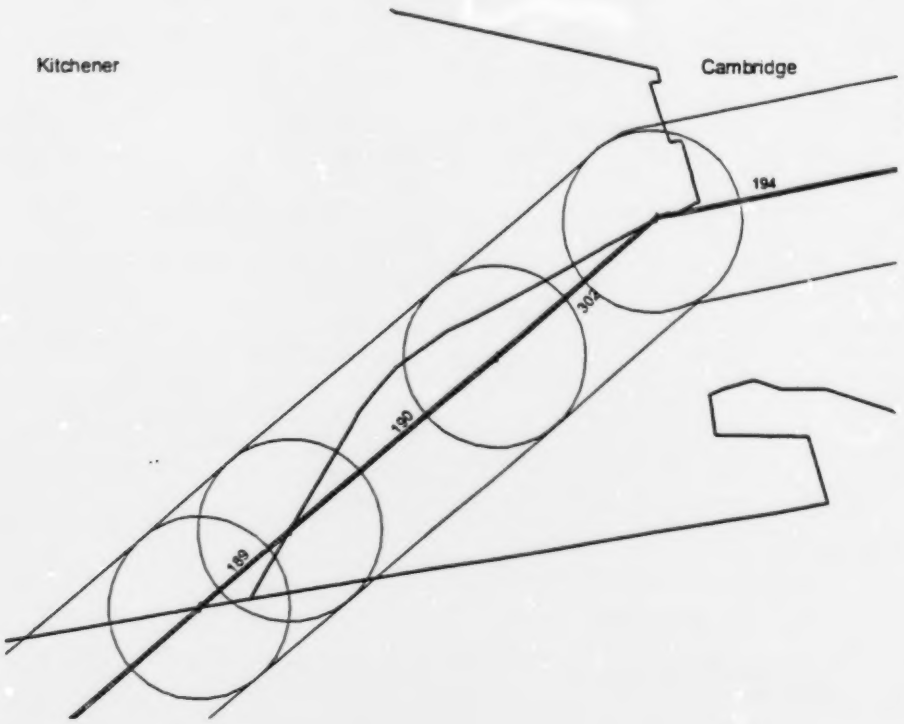

Figure 7: The exposure zones in Kitchener and Cambridge, Ontario

Although there are a number of alternative paths between Halton Hills and Ancaster, two of them dominate the others when a shortest path algorithm is implemented by using population exposure as the arc impedance. These two paths are denoted as Path I (O-A-C-H-D) and Path II (O-A-E-D). The nodes in these routes are depicted in Figure 6. Table 1 shows the details of our calculations.

			Exposed Population within 800 metres	
Links	Path I	Path II	Semicircular	Rectangular
153	*	*	226	93
152	*	*	216	153
151	*	*	100	22
18	*	*	0	0
21	*		52	19
176	*		2092	1752

			Corrections		
Link pairs	Path I	Path II	Semicircular	Rectangular	Angle
153-152	*	*	109		
152-151	*	*	99		
151-18	*	*	0		
21-176	*		48		
176-293	*		401		
293-267	*		859	16	123

293	*		2469	2013
67	*		905	347
178	*		689	268
295	*		575	446
23	*		1655	1512
301	*		888	698
195		*	40	2
194		*	1535	946
302		*	1716	1217
190		*	1957	1131
189		*	1580	858
188		*	1063	509
269		*	0	0
28		*	0	0
184		*	0	0
185		*	0	0
186		*	0	0
187		*	0	0
299		*	129	59
300		*	1609	1519

267-178	*		436		
178-295	*		103	2	119
295-23	*		150		
23-301	*		193	0	162
195-194		*	40		
194-302		*	877	3	151
302-190		*	448		
190-189		*	961		
189-188		*	874		
299-300		*	96		

Table 1: The population exposure calculations

Table 2 summarizes the results of our comparative analysis. When the standard shortest path algorithm is used, Path I is selected with semicircular exposure zones, whereas Path II is selected with rectangular exposure zones. Not only does the use of the semicircular zone result in a very significant overestimation, it also results in the selection of the wrong path. In contrast, the impedance-adjusting shortest path algorithm produces error-free results irrespective of the type of exposure zone used (the difference between the figures are due to roundoff errors).

	Standard Dijkstra		Impedance-adjusted	
	Semicircular	Rectangular	Semicircular	Rectangular
Path I	9867	7485	7469	7467
Path II	10171	6671	6667	6668

Table 2: Summary of the results